A Way Out of the Learning-Rate Morass:
Quantity as an Independent Variable (QAIV)

Stephen A. Book and Erik L. Burgess

Chief Technical Director
MCR, Inc.
(310) 640-0005 x244
sbook@mcri.com

Technical Manager
MCR, Inc.
(703) 633-2128
eburgess@mcri.com

Advanced Training Track
36th Annual DoD Cost Analysis Symposium
Williamsburg, VA
28-31 January 2003

An AT&T Company

Abstract

Standard cost-estimating practice involves application of a cost-improvement factor, or "learning" rate, to account for management, engineering, and production improvements that save money as successive units are produced, although it is difficult or even impossible to determine what the "correct" learning rate will be in any particular estimating context. Unfortunately, the estimator's choice of learning rate exerts a major, perhaps dominant, impact on the estimate of the total spending profile of a large-quantity production program, to the extent that small variations in the assumed learning rate substantially outweigh all other contributions to the total program estimate. Furthermore, sharp disagreements between program offices and organizations overseeing them regarding which learning rate is to be used in estimating costs of multiple-unit procurements are rapidly becoming routine. Some recent disagreements have been especially noteworthy when T1 (first-unit) costs have been estimated using CERs that were derived from a data base normalized to 95% learning, while costs of production are then run down, say, a 90% learning curve. This process causes the estimates to suffer from the well known "double low-ball effect." The increasingly intense nature of these controversies, especially in the age of "cost realism" and concerns over "executability" that we are now entering, leads us to wonder whether the difficulties attributable to the concept of learning may very well exceed its value to the estimating process.

The "Quantity as an Independent Variable" (QAIV) method allows an estimator to circumvent all controversial issues surrounding the learning phenomenon. QAIV calls for estimating the cost of a multiple-unit procurement using a CER that includes "number of units produced" and "prior quantity" as cost drivers, along with the usual weight, thrust, power, etc., technical parameter(s). QAIV CERs estimate average unit cost of a lot of N units directly, rather than estimating the theoretical first-unit cost T1 and then "running" the T1 estimate down a learning curve. Results of this one-step QAIV proof-of-concept study seem to indicate that QAIV-based estimates will have a higher degree of credibility than traditional learning-based estimates, not only because they bypass the controversies associated with learning, but also because QAIV CERs tend to be characterized by lower standard errors and higher R-squared values than do T1-based CERs.

An AT&T Company

Acknowledgments

Research leading to the proof-of-concept results reported here was funded by the Intelligence Community (IC) CAIG under its support contracting arrangement with the TASC/MCR team. The authors are especially grateful for the personal interest and support of the government sponsors of this study, Shishu S. Gupta and Nancy St. Louis of the IC CAIG.

The authors also acknowledge the valuable assistance of Roy E. Smoker of MCR and David S. Colf, now of TASC, who conducted large amounts of data and statistical analysis that led to the conclusions of the QAIV study.

An AT&T Company

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

Example: A Set of Historical Cost Data (FY03$) – Chart 1

(The First 18 of 69 Data Points)

Lot Number	Lot ID	Unit Weight	Lot Size	Average Unit Cost
1	Lot A1	985	37	2937.184
2	Lot A2	985	26	2250.485
3	Lot A3	985	9	2271.142
4	Lot A4	985	69	1745.273
5	Lot A5	985	240	686.028
6	Lot A6	985	180	686.423
7	Lot A7	985	284	522.010
8	Lot A8	985	450	448.385
9	Lot A9	985	432	410.812
10	Lot A10	985	430	398.765
11	Lot A11	985	300	419.569
12	Lot B1	985	15	1886.108
13	Lot B2	985	30	2150.431
14	Lot B3	985	60	1233.518
15	Lot B4	985	132	1220.144
16	Lot B5	985	108	943.088
17	Lot B6	985	265	948.201
18	Lot B7	985	265	788.811

Example: A Set of Historical Cost Data (FY03$) – Chart 2

(The Second 18 of 69 Data Points)

Lot Number	Lot ID	Unit Weight	Lot Size	Average Unit Cost
19	Lot B8	985	265	785.550
20	Lot B9	985	149	811.943
21	Lot B10	985	180	747.886
22	Lot B11	985	195	686.764
23	Lot B12	985	420	432.091
24	Lot C1	510	125	336.666
25	Lot C2	510	390	355.054
26	Lot C3	510	1490	227.871
27	Lot C4	510	1593	195.947
28	Lot C5	510	1560	174.429
29	Lot C6	510	2147	166.896
30	Lot C7	510	1679	164.704
31	Lot C8	510	2527	134.800
32	Lot C9	510	947	174.467
33	Lot D1	190	1200	46.155
34	Lot D2	190	2793	30.797
35	Lot D3	190	2603	28.227
36	Lot D4	190	1682	27.639

An AT&T Company

Example: A Set of Historical Cost Data (FY03$) – Chart 3

(The Third 18 of 69 Data Points)

Lot Number	Lot ID	Unit Weight	Lot Size	Average Unit Cost
37	Lot D5	190	2542	27.301
38	Lot D6	190	784	31.504
39	Lot D7	190	1204	28.778
40	Lot E1	190	65	336.851
41	Lot E2	190	1857	55.258
42	Lot E3	190	1999	46.623
43	Lot E4	190	1535	49.316
44	Lot E5	190	2602	33.489
45	Lot E6	190	3224	28.233
46	Lot E7	190	3461	26.808
47	Lot E8	190	2060	26.408
48	Lot E9	190	3667	22.878
49	Lot E10	190	710	27.797
50	Lot E11	190	788	31.033
51	Lot F1	190	920	70.755
52	Lot F2	190	900	42.316
53	Lot F3	190	1100	40.671
54	Lot F4	190	2487	27.138

Example: A Set of Historical Cost Data (FY03$) – Chart 4

(The Final 15 of 69 Data Points)

Lot Number	Lot ID	Unit Weight	Lot Size	Average Unit Cost
55	Lot G1	190	1534	70.673
56	Lot G2	190	1020	61.789
57	Lot G3	190	2000	38.329
58	Lot G4	190	2245	29.667
59	Lot G5	190	2014	34.318
60	Lot H1	510	65	709.236
61	Lot H2	510	29	746.029
62	Lot H3	510	100	840.908
63	Lot H4	510	225	425.016
64	Lot H5	510	600	216.052
65	Lot H6	510	880	204.019
66	Lot H7	510	1110	161.109
67	Lot H8	510	1398	139.891
68	Lot H9	510	900	126.868
69	Lot H10	510	1144	111.023

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

An AT&T Company

Costs of Multiple-Unit Procurements are Typically Estimated in Two Parts

- **1. Nonrecurring Development Cost, the Data Base for which Is not Impacted by Quantity Normalization**

- **2. Recurring Production Cost, which is Usually Initially Expressed in Terms of a "Theoretical" First-Unit Cost (T1), the Data Base for which Is Critically Impacted by Quantity Normalization**
    - **T1-Cost "Data" are Decisively Influenced by Modeler's Choice of Learning* Rate for Normalization**
    - **Therefore T1 CERs, which are Derived from Those Normalized Data, are also Decisively Influenced by Modeler's Choice of Learning Rate**

* We consider "learning" and "cost improvement" to be synonyms.

An AT&T Company

Vocabulary of Learning*

- AUC(*N*)   =   Average Unit Cost (AUC) of *N* Units
          =   (Total Cost of *N* Units) ¸ *N*

- T1 = Theoretical-First-Unit Cost
          = AUC(1)

- Cumulative Average Learning Rate (LR):
    $AUC(2N) = LR \times AUC(N)$

- Learning Curve (LC): Graph of Exponential Algebraic Expression
    $AUC(N) = T1 \times N^{(\log LR)/\log 2}$

* All computations are done here in terms of "cumulative average" learning, rather than "unit" learning. It is not difficult to transform one version of learning into the other, if necessary. The primary modern reference on the mathematics of the learning phenomenon is the 1997 monograph by Dr. David A. Lee, *The Cost Analyst's Companion*.

An AT&T Company

Typical T1 CER-Development Process

- **Collect, Group and Normalize Historical Cost Data From Multiple Programs of Same Kind**
  - **Use Inflation Indices to Normalize Costs to Same Fiscal Year**
  - **Assume Learning Rate to Normalize Quantities (here 84% learning, for example)***
- **For Each Program in Data Base, Use 84% Learning Rate to "Back Out" T1 Value, _viz._**

$$T1 = AUC(N) \cdot N^{(\log LR)/\log 2}$$

- **Regress the T1 Values against Weights to Derive a "Recurring Production" CER that Estimates T1**

* Assumptions of this kind are usually based on studies showing that learning rates for a mix of items of the kind represented by one's sample data range from 75.5% to 95.8% (in our case) over several dozen production lots. The median learning rate for the kinds of items represented here is approximately 84%.

An AT&T Company

Example: Computations to Prepare Data Set for Calculating T1 Values – Chart 1
(The First 18 of 69 Data Points)

Lot ID	Lot Size	Average Unit Cost	Cumulative Lot Size	Lot-Average-Unit Number*	Cumulative Program Cost
Lot A1	37	2,937.18	37	18.5	108,675.81
Lot A2	26	2,250.49	63	50.5	167,188.42
Lot A3	9	2,271.14	72	68.0	187,628.70
Lot A4	69	1,745.27	141	107.0	308,052.53
Lot A5	240	686.03	381	261.5	472,699.29
Lot A6	180	686.42	561	471.5	596,255.44
Lot A7	284	522.01	845	703.5	744,506.18
Lot A8	450	448.38	1295	1070.5	946,279.28
Lot A9	432	410.81	1727	1511.5	1,123,750.23
Lot A10	430	398.77	2157	1942.5	1,295,219.23
Lot A11	300	419.57	2457	2307.5	1,421,089.79
Lot B1	15	1,886.11	15	7.5	28,291.63
Lot B2	30	2,150.43	45	30.5	92,804.55
Lot B3	60	1,233.52	105	75.5	166,815.65
Lot B4	132	1,220.14	237	171.5	327,874.69
Lot B5	108	943.09	345	291.5	429,728.24
Lot B6	265	948.20	610	478.0	681,001.39
Lot B7	265	788.81	875	743.0	890,036.27

* Includes consideration of prior quantity, as well as current lot quantity.

Example: Computations to Prepare Data Set for Calculating T1 Values – Chart 2
(The Second 18 of 69 Data Points)

Lot ID	Lot Size	Average Unit Cost	Cumulative Lot Size	Lot-Average-Unit Number*	Cumulative Program Cost
Lot B8	265	785.55	1140	1008.0	1,098,206.96
Lot B9	149	811.94	1289	1215.0	1,219,186.44
Lot B10	180	747.89	1469	1379.5	1,353,805.91
Lot B11	195	686.76	1664	1567.0	1,487,724.86
Lot B12	420	432.09	2084	1874.5	1,669,202.88
Lot C1	125	336.67	125	62.5	42,083.31
Lot C2	390	355.05	515	320.5	180,554.40
Lot C3	1490	227.87	2005	1260.5	520,082.36
Lot C4	1593	195.95	3598	2802.0	832,225.73
Lot C5	1560	174.43	5158	4378.5	1,104,335.34
Lot C6	2147	166.90	7305	6232.0	1,462,660.31
Lot C7	1679	164.70	8984	8145.0	1,739,197.93
Lot C8	2527	134.80	11511	10248.0	2,079,838.66
Lot C9	947	174.47	12458	11985.0	2,245,059.09
Lot D1	1200	46.15	1200	600.0	55,385.67
Lot D2	2793	30.80	3993	2597.0	141,400.30
Lot D3	2603	28.23	6596	5295.0	214,874.77
Lot D4	1682	27.64	8278	7437.5	261,364.15

* Includes consideration of prior quantity, as well as current lot quantity.

Example: Computations to Prepare Data Set for Calculating T1 Values – Chart 3

(The Third 18 of 69 Data Points)

Lot ID	Lot Size	Average Unit Cost	Cumulative Lot Size	Lot-Average-Unit Number*	Cumulative Program Cost
Lot D5	2542	27.30	10820	9549.5	330,762.36
Lot D6	784	31.50	11604	11212.5	355,461.69
Lot D7	1204	28.78	12808	12206.5	390,110.04
Lot E1	65	336.85	65	32.5	21,895.30
Lot E2	1857	55.26	1922	994.0	124,509.61
Lot E3	1999	46.62	3921	2922.0	217,708.88
Lot E4	1535	49.32	5456	4689.0	293,409.29
Lot E5	2602	33.49	8058	6757.5	380,546.90
Lot E6	3224	28.23	11282	9670.5	471,569.77
Lot E7	3461	26.81	14743	13013.0	564,351.37
Lot E8	2060	26.41	16803	15773.5	618,751.64
Lot E9	3667	22.88	20470	18637.0	702,646.75
Lot E10	710	27.80	21180	20825.5	722,382.50
Lot E11	788	31.03	21968	21574.5	746,836.17
Lot F1	920	70.76	920	460.0	65,094.78
Lot F2	900	42.32	1820	1370.5	103,179.09
Lot F3	1100	40.67	2920	2370.5	147,917.60
Lot F4	2487	27.14	5407	4164.0	215,409.26

* Includes consideration of prior quantity, as well as current lot quantity.

Example: Computations to Prepare Data Set for Calculating T1 Values – Chart 4

(The Final 15 of 69 Data Points)

Lot ID	Lot Size	Average Unit Cost	Cumulative Lot Size	Lot-Average-Unit Number*	Cumulative Program Cost
Lot G1	1534	70.67	1534	767.0	108,411.88
Lot G2	1020	61.79	2554	2044.5	171,436.20
Lot G3	2000	38.33	4554	3554.5	248,094.60
Lot G4	2245	29.67	6799	5677.0	314,697.48
Lot G5	2014	34.32	8813	7806.5	383,814.14
Lot H1	65	709.24	65	32.5	46,100.36
Lot H2	29	746.03	94	80.0	67,735.20
Lot H3	100	840.91	194	144.5	151,826.02
Lot H4	225	425.02	419	307.0	247,454.59
Lot H5	600	216.05	1019	719.5	377,085.65
Lot H6	880	204.02	1899	1459.5	556,622.58
Lot H7	1110	161.11	3009	2454.5	735,453.30
Lot H8	1398	139.89	4407	3708.5	931,021.03
Lot H9	900	126.87	5307	4857.5	1,045,202.33
Lot H10	1144	111.02	6451	5879.5	1,172,212.72

* Includes consideration of prior quantity, as well as current lot quantity.

A Typical Method of "Backing-Out" a T1 Value for Each Program

- **For Each Program in the Data Base, Use an 84% Learning Rate to "Back Out" a T1 Value, *viz.***

$$T1 = AUC(N) \div N^{(\log LR)/\log 2}$$

- **For Example, Consider Program A, which Included 2,457 Units and Had an Average Unit Cost of $578.384*, so that**

$$T1 = \$578.384 \div 2,457^{(\log 0.84)/\log 2} = \$4,142.304$$

- **The Next Charts List the Backed-Out T1 Values for All Eight Programs on which Regression Will be Done in Order to Derive a T1 CER**

* **Calculated by dividing the cumulative (total) program cost by the cumulative lot size (total number of units).**

Example: "Backed-Out" T1 Values for All Eight Programs

Program	Unit Weight	Production Run Total Units	Production Run Total Cost	Average Unit Cost	Backed-Out T1 (84% Learning)
A	985	2,457	1,421,089.787	578.384	4,121.304
B	985	2,084	1,669,202.883	800.961	5,475.740
C	510	12,458	2,245,059.095	180.210	1,931.717
D	190	12,808	390,110.040	30.458	328.773
E	190	21,968	746,836.172	33.997	420.304
F	190	5,407	215,409.259	39.839	346.171
G	190	8,813	383,814.141	43.551	427.906
H	510	6,451	1,172,212.716	181.710	1,650.624

An AT&T Company

A Weight-Based CER that Estimates T1

- **Regress T1 Values against Unit Weights to Derive a "Recurring Production" CER that Estimates T1 for Items Similar to those Represented in the Data Base**

- **A Good Multiplicative-Error* CER Appears to be the Following:**

$$T1 = 27.24 + 0.12\ W^{1.54}$$

- **CER Quality Statistics**
  - **Standard Error of the Estimate = 14.53%**
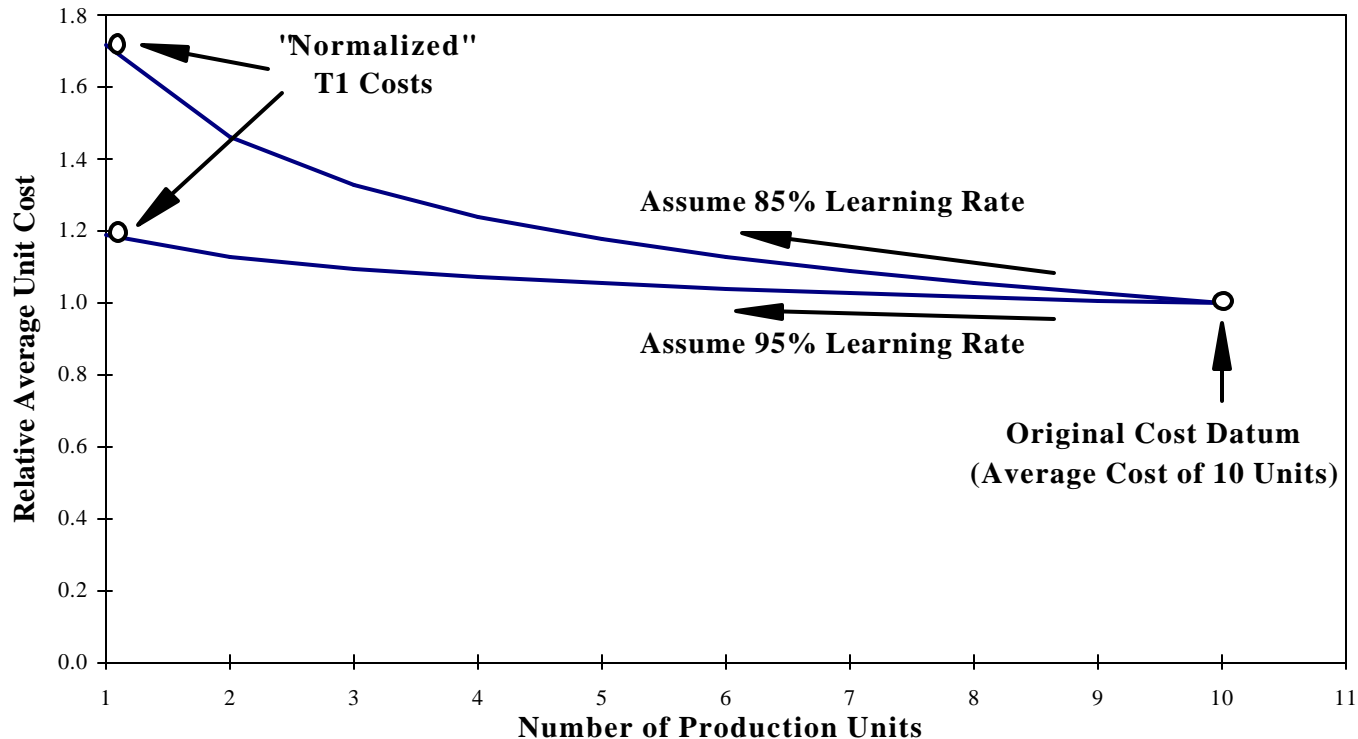 - **Pearson's R^2 = 96.41%****
  - **Percentage Bias = 0.00%**

*   Reference: S.A. Book and N.Y. Lao (1998).
** R^2 between CER-based estimates and data base "actuals."

An AT&T Company

Data Base for Weight-Based CER for T1

T1 vs. Weight
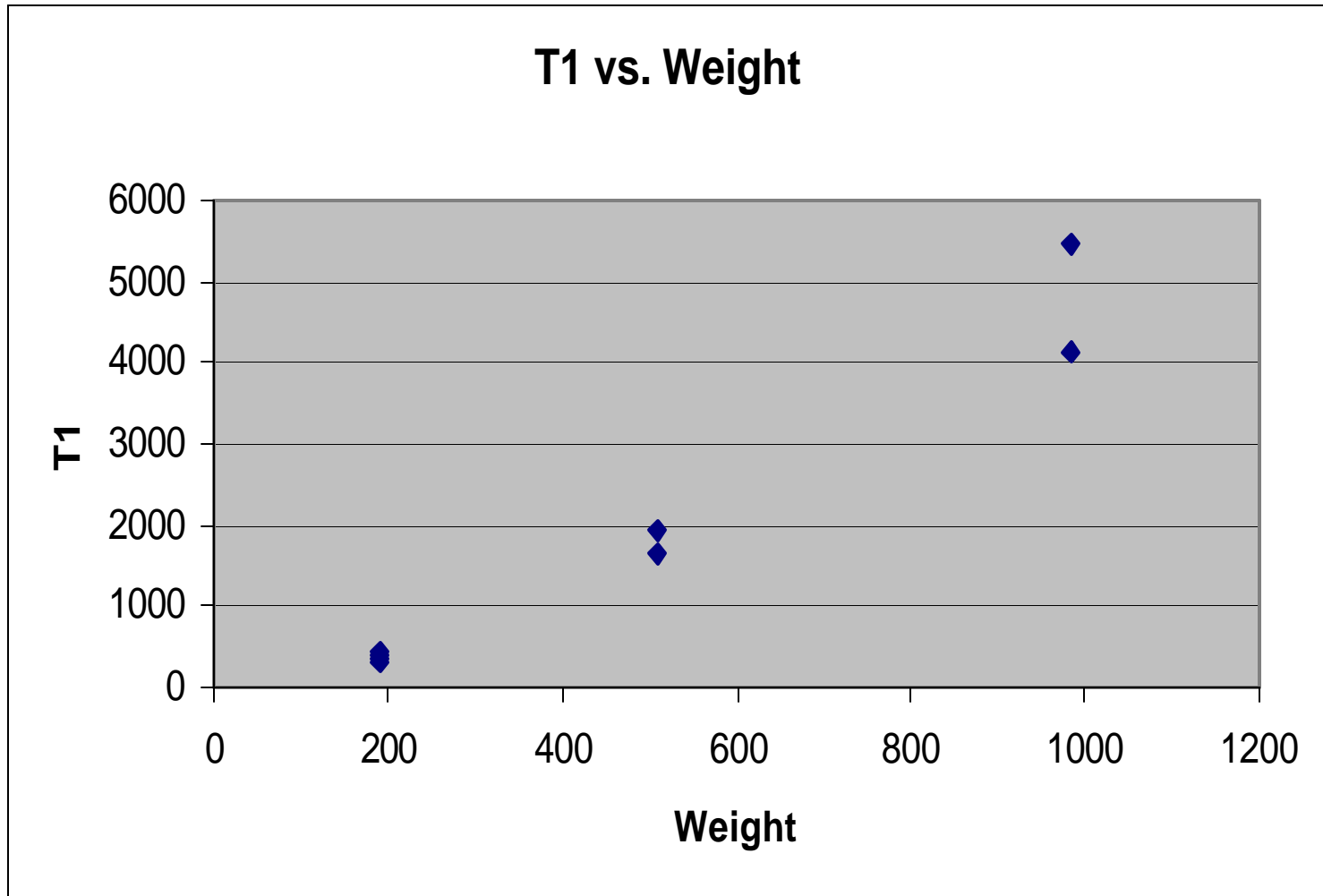
A Sad Fact about Data Normalization: Assumed Learning Rate Impacts T1

- **Program Cost Is Supplied to the Analyst as, for example, AUC(10)**
- **Costs are Normalized to T1 Using 85% or 95% Learning, Respectively**
- **Normalized T1 Is 45% Higher at 85% Learning Than at 95% Learning**
- **So the T1 Values Supporting the CER are not Really "Actuals" (If they were "actuals", we would refer to the first-unit cost as A1 instead of T1!)**

Reference: S.A. Book and E.L. Burgess (1996).

A "More Logical" CER-Development Process for T1

- **Collect, Group and Normalize* Historical Cost Data From Several Programs of the Same Kind**
  - Use Inflation Indices to Normalize Costs to Same FY
  - Use All Lots of Each Program to Fit a Learning Curve to that Program's Lot Data
  - Regress Lot-Average-Unit Cost (AUC) against Lot-Average-Unit Number (see Charts #12-15)
  - Derive an "Actual" T1 and Learning Rate for Each Program

- **For Each Program in Data Base, We Will Have a T1 Value that is Characteristic of that Program and Does Not Require the Same Learning Rate to be Assumed for All Programs**

- **Regress T1 Values against Weights to Derive a "Recurring Production" CER that Estimates T1**

* Normalize then-year data to base year using inflation effects only – no quantity normalization.

An AT&T Company

Example: Calculated T1 Values and Learning Rates for All Eight Programs

Program	Unit Weight	Number of Lots	Learning Rate	T1	Standard Error	Pearson's R-squared	Percentage Bias
A	985	11	71.90%	13,724	17.65%	93.60%	0.00%
B	985	12	84.97%	3,788	16.52%	82.78%	0.00%
C	510	9	88.20%	834	12.51%	84.51%	0.00%
D	190	7	90.26%	110	10.21%	80.63%	0.00%
E	190	11	76.54%	1,113	18.21%	98.72%	0.00%
F	190	4	75.61%	834	10.64%	96.90%	0.00%
G	190	5	64.44%	13,456	22.21%	76.87%	0.00%
H	510	10	72.98%	5,576	22.54%	75.92%	0.00%

A More Logically-Derived Weight-Based CER that Estimates T1

- **Regress T1 Values against Unit Weights to Derive a "Recurring Production" CER to Estimate T1 for Items Similar to those Represented in Data Base**

- **The Best Multiplicative-Error* CER Appears to be the Following:**

$$T1 = 834.00 + 48.57\ W^{0.78}$$

- **CER Quality Statistics**
  - **Standard Error of the Estimate = 182.24%**
  - **Pearson's R^2 = 13.16%****
  - **Percentage Bias = 0.00%**

- **Now We Know One Reason CER Developers Like to Normalize all Programs to Same Learning Rate!**

* **Reference: S.A. Book and N.Y. Lao (1998).**
** **R^2 between CER-based estimates and data base "actuals."**

Data Base for Logically-Derived Weight-Based CER for T1
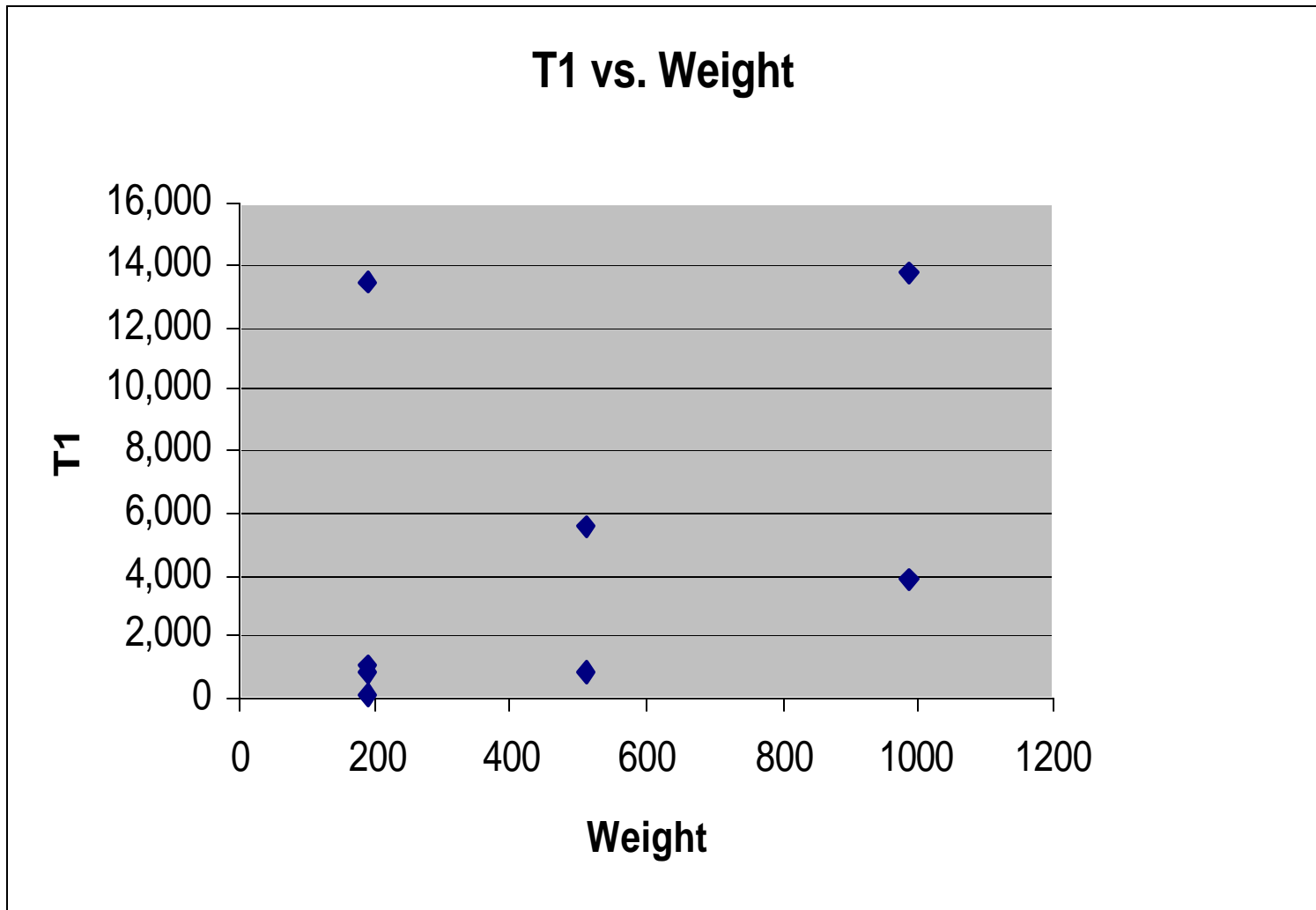
T1 vs. Weight

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

Estimating Average Unit Cost of a Multiple-Unit Production Run

- **Suppose We Are Tasked to Estimate the Average Unit Cost of a Lot of 100 Units of a Similar Item that Weighs 300 Pounds**
- **We Will Estimate This Cost under Three Different Conditions on Where the Lot is Located in the Item's Entire Production History**
  - **Just After the 10th Unit**
  - **Just After the 200th Unit**
  - **Just After the 3000th Unit**

An AT&T Company

The First Step is to Estimate T1

- **Using Our First T1 CER, We Estimate T1 as Follows:**

$$T1 = 27.24 + 0.12 \; W^{1.54}$$
$$= 27.24 + 0.12 \, (300)^{1.54} = 810.58$$

- **The Next Step is to Choose a Learning Rate and "Run the T1 Estimate Down the Learning Curve"**

- **Then, Starting with the $(Q+1)^{st}$ Unit, the AUC of 100 Units is**

$$AUC(100) = \frac{(Q+100)AUC(Q+100) - Q \; AUC(Q)}{100}$$

"Running It Down the Learning Curve"

- **Using the Same Learning Rate as Earlier (84%), We Compute the Average Unit Cost of 10, 110, 200, 300, 3000, and 3100 Units**

- **(log 0.84)/(log 2) = - 0.251538767**

- **Therefore**

$$AUC(10) = 810.58 \times 10^{(log\ 0.84)/log2} = 454.21$$

$$AUC(110) = 810.58 \times 110^{(log\ 0.84)/log2} = 248.49$$

$$AUC(200) = 810.58 \times 200^{(log\ 0.84)/log2} = 213.80$$

$$AUC(300) = 810.58 \times 300^{(log\ 0.84)/log2} = 193.07$$

$$AUC(3000) = 810.58 \times 3000^{(log\ 0.84)/log2} = 108.18$$

$$AUC(3100) = 810.58 \times 3100^{(log\ 0.84)/log2} = 107.30$$

Now to Finish Our Estimating Task …

- **Just After the 10th Unit**

$$\text{AUC(100)} = \frac{(110)\,\text{AUC(110)} - 10\,\text{AUC(10)}}{100}$$

$$= \frac{(110)(248.49) - 10(454.21)}{100} = 227.92$$

- **Similarly, Just After the 200th Unit**

$$\text{AUC(100)} = \frac{(300)(193.07) - 200(213.80)}{100} = 151.61$$

- **Just After the 3000th Unit**

$$\text{AUC(100)} = \frac{(3100)(107.30) - 3000(108.18)}{100} = 80.90$$

An AT&T Company

But What if 84% is <u>Not</u> the Correct Learning Rate?

- **Recall that We Chose 84% from a Set of Learning Rates Actually Experienced that Ranged from 75.5% to 95.8%**
- **For Estimating AUC of 100 Units Following the 10th Unit, This Means that AUC(110) Could Credibly Range from**

$$AUC(110) = 810.58 \times 110^{(\log 0.755)/\log 2} = 120.53$$
$$\text{to } AUC(110) = 810.58 \times 110^{(\log 0.958)/\log 2} = 605.94$$

- **Percentagewise, this Range Represents a Possible Decrease in Average Unit (equivalently, Total) Cost of 51.50% up to a Possible Increase of 143.85%**
- **These Possible Percentage Errors will Carry Over to the Estimate of AUC(100)**

Estimated AUC(100) at Three Different Learning Rates

AUC(100) Just After the …	84% Learning	90% Learning	95% Learning
10th Unit	227.92	379.28	561.33
200th Unit	151.61	297.32	499.10
3000th Unit	80.90	202.83	414.42

Note: These Numbers are Calculated by the Same Process as that on Charts #28-29.

Learning Assumptions Impact Cost Estimates
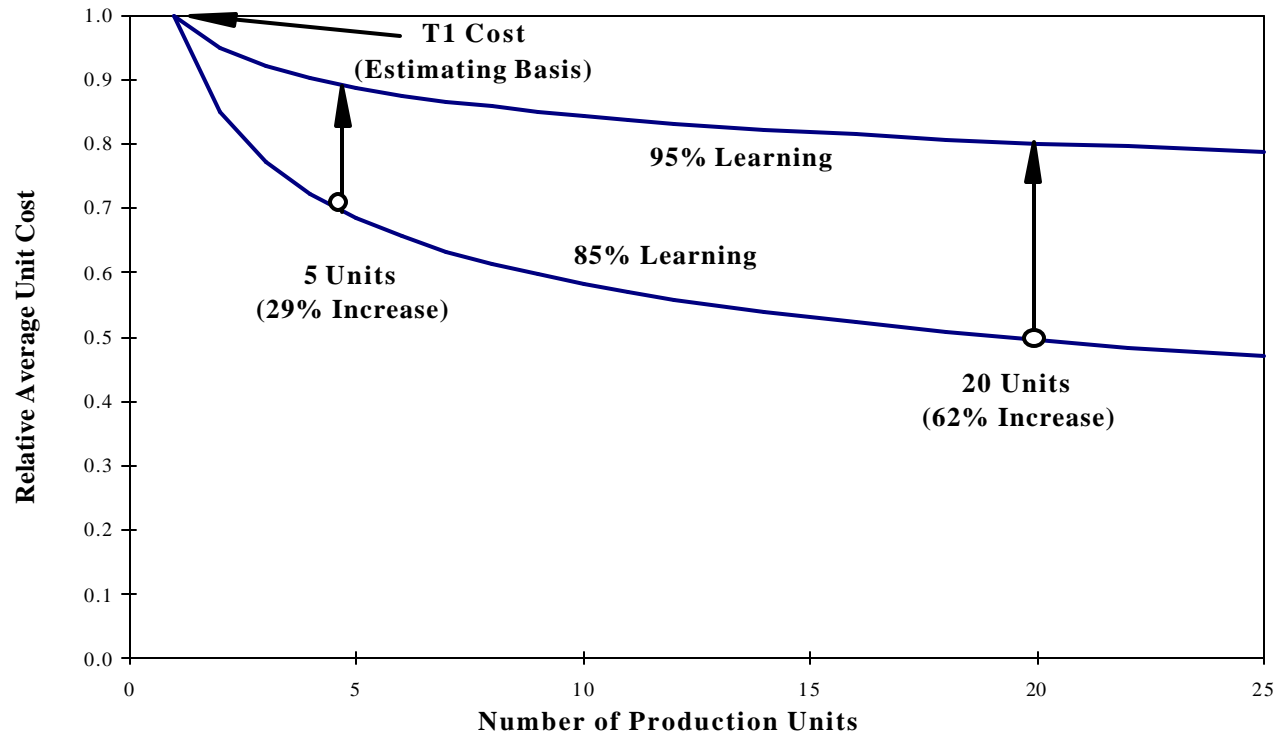
- **All Learning Curves for an Estimate Intersect at Estimated T1 Point**
- **Costs of Multiple Units are Estimated at 85% or 95% Learning**
- **Estimate for 20 units is 62% Higher at 95% Learning than at 85%**

Reference: S.A. Book and E.L. Burgess (1996).

AUC Ratios: *x*% Learning vs. *y*% Learning

Learning Rates	Number Of Units Produced (N)						
(*x*% to *y*%)	10	20	50	100	200	500	1000
100% to 95%	1.19	1.25	1.34	1.41	1.48	1.58	1.67
95% to 90%	1.20	1.26	1.36	1.43	1.51	1.62	1.71
90% to 85%	1.21	1.28	1.38	1.46	1.55	1.67	1.77
85% to 80%	1.22	1.30	1.41	1.50	1.59	1.72	1.83
80% to 75%	1.24	1.32	1.44	1.54	1.64	1.78	1.90
75% to 70%	1.26	1.35	1.48	1.58	1.69	1.86	1.99

To use this table to calculate other ratios (without going back to the mathematical formula), multiply successive numbers in the appropriate column. For example, the AUC Ratio for 95% to 85% learning when the production run consists of 10 units is 1.20 ´ 1.21 = 1.45, the product of the ratios for 95% to 90% and 90% to 85%, respectively. The AUC Ratio for 90% to 80% learning when the production run consists of 500 units is 1.67 ´ 1.72 = 2.87, the product of the ratios for 90% to 85% and 85% to 80%, respectively. The ratio of 2.87 can be interpreted as follows: If we assume a learning rate of 90% to estimate costs of a production run of 500 units, our cost estimate will be 187% higher than if we had assumed a learning rate of 80%. **In other words, if we assume a learning rate of 80% for our estimate, but the actual learning rate turns out to be 90%, we will experience a 187% overrun in production cost.** (This chart is based on "cumulative-average" learning theory.)

Reference: S.A. Book and E.L. Burgess (1996).

Combined T1 and Learning Errors of Estimation

- **Recall that the Percentage Standard Error ("one sigma") of the Estimate for the Original T1 CER (on Chart #18) is 14.53%**

- **Assuming the Range -51.50% to +143.85% Serves as a "Three-Sigma" (Nonsymmetric) Interval around the AUC(100) Estimate, an "Average" One-Sigma Value for the Learning Effect is [(51.5%¸3)+(143.85%¸3)]¸2 = 32.56%**

- **Combining These Two Sigma Values (One for the T1 CER and the Other for Learning) by Root-Sum-Square (assuming independence of the estimating and learning effects), We Obtain an "Average" Standard Error of 35.65% in Our Estimate of AUC(100)**

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

Some Serious Problems* with Learning Theory

- **Learning Theory Requires that Cost Estimating Be Done in a Two-Step Process**
  - We First Estimate T1 as a Function of a Cost Driver (e.g., weight), which is a Historical Fact
  - Then We Choose a Learning Rate, Often Based on Historical Experience but Sometimes Not
- **The Learning Rate (even for a completed program) is Rarely a Historical Fact, but rather Something that Must be Estimated from Experience – It is also Notoriously Difficult to Forecast in Advance and Probably Introduces Bias into the Estimate**
- **Each of these Estimating Steps Introduces an Error of Some Magnitude, and the Combination of the Two Errors is Likely to be Very Large**

•Additional serious difficulties of a more sophisticated nature were recently identified by T.P. Anderson (2002) and J.C. Latta (2002).

An AT&T Company

QAIV Theory

- **QAIV Theory Calls for Estimating to be Done in One Step**
  - **One Way to Do This is to Use a Bivariate CER Having both Weight and Lot Size* as Cost Drivers**
  - **This CER Can be Based Entirely on Historical Data, as both Weight and Lot Size are Known Historical Facts**
  - **No Learning-Rate Assumption is Needed for either Data-Base Normalization or Estimating**
  - **Furthermore, This One-Step Estimating Process Has Only One Standard Error of the Estimate**

- **We Regress the Average-Unit-Cost Values against Weights and Lot Sizes to Derive a "Recurring Production" CER that Estimates AUC for Multiple-Unit Production Lots**

* This idea apparently originated in work of D. MacKenzie and his colleagues at Wyle Laboratories in the early 1990s. Later in the decade, they looked at the efficacy of CERs that estimate AUC per Pound. A recent discussion (from another direction) of the idea of including "lot size" as a cost driver in the context of CER-based estimating can be found in a recent report by S.A. Book and J.C. Latta (2000).

An AT&T Company

Comparative Benefits of QAIV-Based Estimates

Are these Sources of Estimating Uncertainty …	when Using T1 CERs?	when Using QAIV CERs?
1. Unknown true learning rate that requires modeler to pick one to "back out" T1 value, thereby causing unreported CER bias.	Yes	No
2. Normalizing all T1 values to one rate, when the rates in fact differ, that distorts data base and causes unreported CER standard error.	Yes	No
3. Regressing T1s against a single cost driver, such as weight, which ignores impact of other cost drivers.	Yes	Yes
4. Picking learning rate after debate to "run estimate down learning curve", when in fact learning rate of future program is unknown.	Yes	No
5. Technical uncertainties and other risk issues characteristic of program being estimated that introduce additional uncertainty.	Yes	Yes

An AT&T Company

Example: Portion of Data Base Used to Derive Bivariate CER – Chart 1

(The First 36 of 69 Data Points)

Unit Weight	Lot Size	Average Unit Cost	Unit Weight	Lot Size	Average Unit Cost
985	37	2937.184	985	265	785.550
985	26	2250.485	985	149	811.943
985	9	2271.142	985	180	747.886
985	69	1745.273	985	195	686.764
985	240	686.028	985	420	432.091
985	180	686.423	510	125	336.666
985	284	522.010	510	390	355.054
985	450	448.385	510	1490	227.871
985	432	410.812	510	1593	195.947
985	430	398.765	510	1560	174.429
985	300	419.569	510	2147	166.896
985	15	1886.108	510	1679	164.704
985	30	2150.431	510	2527	134.800
985	60	1233.518	510	947	174.467
985	132	1220.144	190	1200	46.155
985	108	943.088	190	2793	30.797
985	265	948.201	190	2603	28.227
985	265	788.811	190	1682	27.639

Example: Portion of Data Base Used to Derive Bivariate CER – Chart 2

(The Remaining 33 of 69 Data Points)

Unit Weight	Lot Size	Average Unit Cost
190	2542	27.301
190	784	31.504
190	1204	28.778
190	65	336.851
190	1857	55.258
190	1999	46.623
190	1535	49.316
190	2602	33.489
190	3224	28.233
190	3461	26.808
190	2060	26.408
190	3667	22.878
190	710	27.797
190	788	31.033
190	920	70.755
190	900	42.316
190	1100	40.671
190	2487	27.138

Unit Weight	Lot Size	Average Unit Cost
190	1534	70.673
190	1020	61.789
190	2000	38.329
190	2245	29.667
190	2014	34.318
510	65	709.236
510	29	746.029
510	100	840.908
510	225	425.016
510	600	216.052
510	880	204.019
510	1110	161.109
510	1398	139.891
510	900	126.868
510	1144	111.023

Bivariate QAIV CER: AUC as a Function of Weight and Lot Size

- **As Before, Suppose We Are Tasked to Estimate the Average Unit Cost of a Lot of 100 Units of a Similar Kind of Item that Weighs 300 Pounds**

- **A Good Multiplicative-Error* CER Appears to be the Following:**

$$AUC(N) = \text{-}20.72 + 10.43 \ W^{0.98} \ N^{-0.45}$$

- **CER Quality Statistics**
  - **Standard Error of the Estimate = 28.93%**
  - **Pearson's R^2 = 86.3%**
  - **Percentage Bias = 0.00%**

- **No Additional Error Need be Added to Account for Learning-Rate Estimation Errors, Because We Do Not Use a Learning-Rate Estimate**

* Reference: S.A. Book and N.Y. Lao (1998).

Estimating AUC(100) Using the Bivariate QAIV CER

- **As Before, Estimate Average Unit Cost of a Lot of 100 Units of an Item that Weighs 300 Pounds**
- **Using Our Bivariate QAIV CER,**

$$AUC(100) = -20.72 + 10.43(300)^{0.98}(100)^{-0.45} = 330.73$$

- **Compare this Estimate with the T1/Learning-Based Estimates Obtained Earlier:**

AUC(100) Just After the …	84% Learning	90% Learning	95% Learning
10th Unit	227.92	379.28	561.33
200th Unit	151.61	297.32	499.10
3000th Unit	80.90	202.83	414.42

- **One Problem: Bivariate CER Cannot Make Use of Information on Where the Lot of 100 Units Occurs in Program's Overall Production History**

Enhancing the QAIV Approach to Account for Prior-Quantity Effects*

- **We Have the Same Estimating Problem as Before: Estimate the Average Unit Cost of a Lot of 100 Units (Similar to Those in the Data Base) that Each Weigh 300 Pounds**
- **However, it Probably Matters Whether the Lot of 100 Units of Item "X" Constitutes Lot "X5" or Lot "X10"**
  - **The 100 Units in Lot X10 Should Cost Less Per Unit than the 100 Units in Lot X5 (We Think)**
  - **This Time, therefore, We Want to Account for the Effects of Prior Quantity**
  - **This Prior-Quantity Effect is Different from a "Rate Effect"**

* This enhancement to the QAIV methodology was proposed by S.S. Gupta.

44

Trivariate QAIV CER: AUC as a Function of Weight, Lot Size, and Prior Quantity

- **This Estimating Problem Can be Solved Using a "Trivariate" CER, namely a CER that Has Weight (*W*), Lot Size (*N*), *and Prior Quantity (Q)* as Cost Drivers**

- **The Trivariate CER Has the Form**

$$AUC(N) = a + bW^c\,N^d\,Q^e$$

- **Technical Note: Computational Difficulties Caused by Attempting to Raise 0 to a Negative Power Make it Necessary to Define "Prior Quantity" of First Lots to be 1 – This Has Essentially No Effect on the CER**

Example: Portion of Data Base Used to Derive Trivariate CER – Chart 1

(The First 36 of 69 Data Points)

Unit Weight	Lot Size	Prior Quantity	Average Unit Cost	Unit Weight	Lot Size	Prior Quantity	Average Unit Cost
985	37	1	2937.184	985	265	875	785.550
985	26	37	2250.485	985	149	1140	811.943
985	9	63	2271.142	985	180	1289	747.886
985	69	72	1745.273	985	195	1469	686.764
985	240	141	686.028	985	420	1664	432.091
985	180	381	686.423	510	125	1	336.666
985	284	561	522.010	510	390	125	355.054
985	450	845	448.385	510	1490	515	227.871
985	432	1295	410.812	510	1593	2005	195.947
985	430	1727	398.765	510	1560	3598	174.429
985	300	2157	419.569	510	2147	5158	166.896
985	15	1	1886.108	510	1679	7305	164.704
985	30	15	2150.431	510	2527	8984	134.800
985	60	45	1233.518	510	947	11511	174.467
985	132	105	1220.144	190	1200	1	46.155
985	108	237	943.088	190	2793	1200	30.797
985	265	345	948.201	190	2603	3993	28.227
985	265	610	788.811	190	1682	6596	27.639

Example: Portion of Data Base Used to Derive Trivariate CER – Chart 2

(The Remaining 33 of 69 Data Points)

Unit Weight	Lot Size	Prior Quantity	Average Unit Cost	Unit Weight	Lot Size	Prior Quantity	Average Unit Cost
190	2542	8278	27.301	190	1534	1	70.673
190	784	10820	31.504	190	1020	1534	61.789
190	1204	11604	28.778	190	2000	2554	38.329
190	65	1	336.851	190	2245	4554	29.667
190	1857	65	55.258	190	2014	6799	34.318
190	1999	1922	46.623	510	65	1	709.236
190	1535	3921	49.316	510	29	65	746.029
190	2602	5456	33.489	510	100	94	840.908
190	3224	8058	28.233	510	225	194	425.016
190	3461	11282	26.808	510	600	419	216.052
190	2060	14743	26.408	510	880	1019	204.019
190	3667	16803	22.878	510	1110	1899	161.109
190	710	20470	27.797	510	1398	3009	139.891
190	788	21180	31.033	510	900	4407	126.868
190	920	1	70.755	510	1144	5307	111.023
190	900	920	42.316				
190	1100	1820	40.671				
190	2487	2920	27.138				

An AT&T Company

A QAIV CER that Accounts for Prior Quantity

- **Suppose We Face the Same Estimating Problem as Before: Estimate the Average Unit Cost (AUC) of a Lot of 100 Units of a Similar Kind of Item that Weighs 300 Pounds, *Except …***
- **Now We Can *Directly* Estimate the Lot's AUC, Based on How Far Along in the Production Schedule the Lot is Located**
- **Inputs to the CER are Now …**
  - *W* = Weight
  - *N* = Lot Size (Number of Units in This Lot)
  - *Q* = Prior Quantity (Total Number of Units in All Prior Lots)

The Trivariate CER and Its Quality Statistics

- **A Good Multiplicative-Error* CER in This Context Appears to be the Following:**

$$AUC(N) = -11.20 + 1.90\ W^{1.20}\ N^{-0.33}\ Q^{-0.08}$$

- **CER Quality Statistics**
  - **Standard Error of the Estimate = 24.24%**
  - **Pearson's R^2 = 90.0%**
  - **Percentage Bias = 0.00%**

- **In Addition to its Consideration of Prior Quantity, There is Noticeable Improvement Relative to the Bivariate CER in Standard Error of the Estimate and Pearson's R^2**

* Reference: S.A. Book and N.Y. Lao (1998).

An AT&T Company

Estimating AUC(100) Using the Trivariate QAIV CER

- **Same Estimating Problem: Estimate Average Unit Cost of a Lot of 100 Units of an Item that Weighs 300 Pounds**
- **Our Trivariate QAIV CER is:**

$$AUC(100) = -11.20 + 1.90(300)^{1.20}(100)^{-0.33}Q^{-0.08}$$

- **This CER Gives Us More Flexibility – We Now Can Directly Estimate Lot AUC Assuming …**
  - **A Prior Quantity of Q = 10 Units:**

    $$AUC(100) = -11.20 + 1.90(300)^{1.20}(100)^{-0.33}(10)^{-0.08} = 313.36$$
  - **A Prior Quantity of Q = 200 Units:**

    $$AUC(100) = -11.20 + 1.90(300)^{1.20}(100)^{-0.33}(200)^{-0.08} = 244.20$$
  - **A Prior Quantity of Q = 3000 Units:**

    $$AUC(100) = -11.20 + 1.90(300)^{1.20}(100)^{-0.33}(3000)^{-0.08} = 194.45$$
- **No Unsupportable, and therefore Controversial, Learning Assumption is Required**

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

Traditional Learning Approach

- **Weight-Based CER for T1 …**
  - **… Has Some Percentage Standard Error ("one-sigma") of the Estimate, Usually in the Range 10%-60%**
  - **… which is Probably too Low, because the Regressed T1 Values are Not Actuals, due to Both Quantity and Inflation Normalization**
- **Additional "Average" One-sigma Error Induced by Learning Curve Amounts to about 15% for Each ±5% Change in the Learning Rate (Based on Our Calculated 32.56% for the Range 75% to 95% Learning)**
- **Combining These Two Sigma Values by Root-Sum-Square (assuming independence of the estimating and learning effects), We Obtain a Standard Error of around 35% Under "Average" Assumptions**

An AT&T Company

QAIV Approach

- **Bivariate and Trivariate QAIV CERs for AUC(N) Generally Will Have …**
  - **… Smaller Standard Error than One-Driver T1 CER due to Additional Explanatory Variable(s)**
  - **… No Error Source other than Standard Error because Regressed AUC(N) Values are Actuals (Except for Inflation Adjustments) - Learning in Data Base Normalization or in Recurring-Cost Estimation**
- **All CER Quality Statistics Will Improve**
  - **Standard Error of the Estimate Will Decrease**
  - **Pearson's R^2 Will Increase**
  - **Percentage Bias Will Remain at 0.00%**

Comparative Estimate Summary

Estimating Method	AUC(100) Just After the …		
	10th Unit	200th Unit	3000th Unit
T1 + 84%LR	227.92	151.61	80.90
T1 + 90%LR	379.28	297.32	202.83
T1 + 95%LR	561.33	499.10	414.42
QAIV(*W*,*N*)	330.73	330.73	330.73
QAIV(*W*,*N*,*Q*)	313.36	244.20	194.45

T1 = Theoretical First-Unit Cost
LR = Learning Rate
W = Weight
N = Lot Size
Q = Prior Quantity

An AT&T Company

Does the QAIV Estimate Support a Learning-Type Phenomenon?

QAIV vs. Learning

Legend:
- LR = 87%
- LR = 88%
- LR = 89%
- LR = 90%
- LR = 91%
- QAIV Trivariate

Y-axis: AUC(100) after LAUN
X-axis: Lot-Average-Unit Number (LAUN)
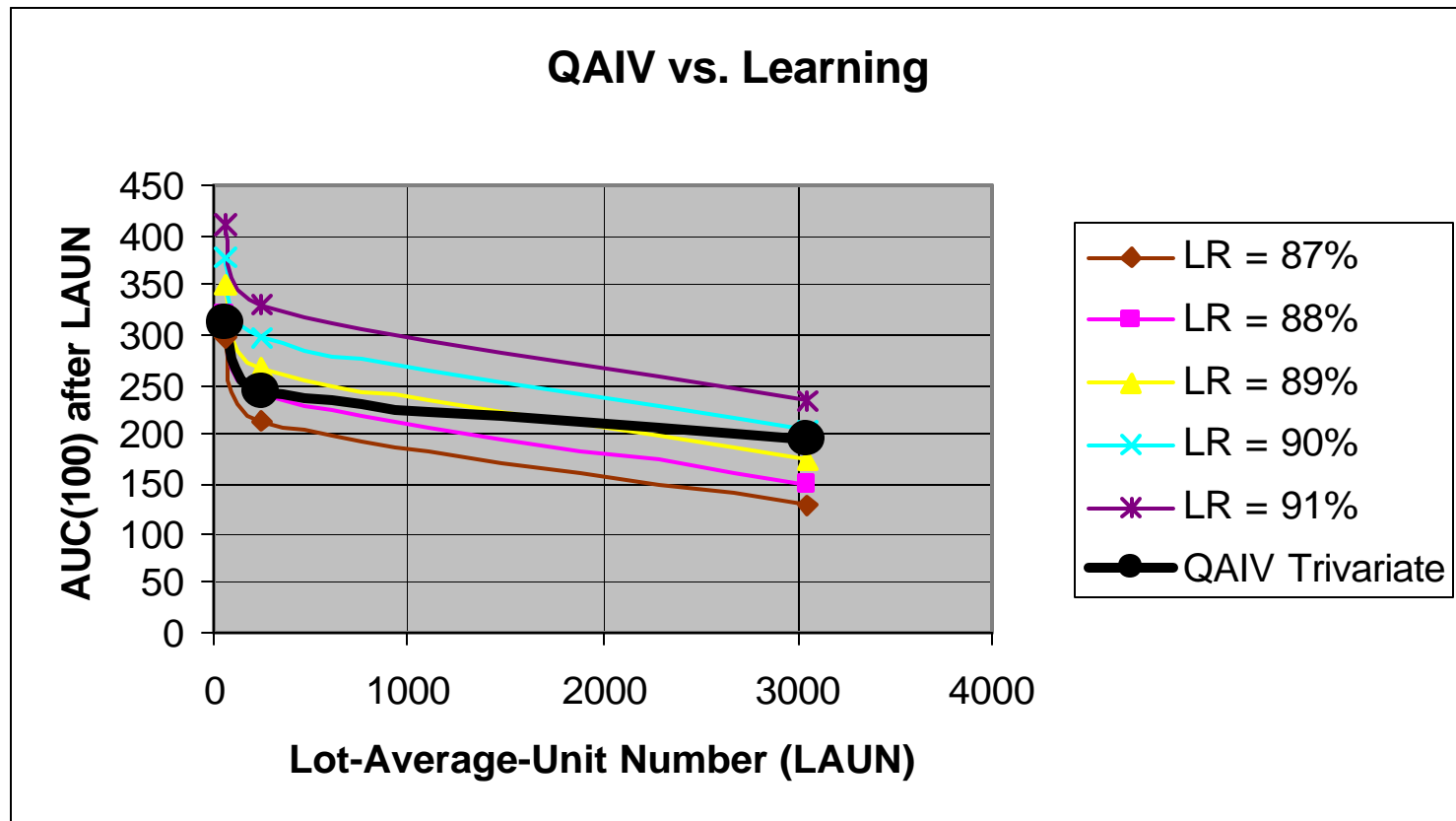
Apparently Not – The QAIV Trivariate Trend Line Does Not Appear to be Exponential in Nature

An AT&T Company

Lessons Learned from Comparison

- **Learning Rate and T1 Values are Not Found in Historical Cost Data Bases**
  - Normalization for Learning Exerts a Strong Impact on Estimates Based on CERs Derived from that Data Base
  - Normalizing all Data Points to Same Learning Rate Appears to be Logically Unsupportable, though Very Convenient
- **Bivariate QAIV CER Appears to Forecast AUC of <u>Initial</u> Lot at a "Realistic", though Unstated, Learning Rate**
- **Trivariate QAIV CER, Using Combination of Cost Driver, Lot Size, and Prior Quantity, Better Models Actual Content of Historical Data Base**
  - QAIV CERs are Not Derived from a Learning-Normalized Data Base
  - No Unsupportable, and therefore Controversial, Learning Assumption is Required

An AT&T Company

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

Summary

- **Traditional Recurring-Production-Cost Estimating Process Consists of Two Steps, Each of which Introduces A Significant Amount of Error**
  - **Step 1: Estimate "Theoretical" First-Unit Cost T1, using CER whose Data Base Must be Normalized not only for Inflation, but also for Quantity according to a Vaguely Defined, and Often Controversial, Learning Rate**
  - **Step 2: Run the T1 Estimate Down a Learning Curve, the Correct Rate for which is Notoriously Difficult to Forecast Accurately**

- **QAIV Process Has Only One Step**
  - **Using CER that Estimates AUC, whose Data Base Undergoes Only One Normalization, that for Inflation**
  - **QAIV Process Not Only Results in Tighter Standard Error than Does the Traditional Learning-Based Approach, but Does Not Contain Additional "Hidden" Errors Associated with Imputed Learning Rates**

An AT&T Company

Conclusion

- **This Proof-of-Concept Study Has Revealed Probable Benefits of QAIV vs. the Traditional Learning-Based Approach**
- **QAIV Appears to be Worth Serious Consideration as an Estimating Methodology for Costs of Multiple-Unit Procurements**

Contents

- **Estimating Costs of Multiple-Unit Procurements**
- **The Traditional "Learning" Approach and Its Discontents**
- **The QAIV Approach**
- **Comparison Between the Two Approaches**
- **Summary and Conclusion**
- **References**

References

- T.P. Anderson, "The Trouble with Learning Curves," Space Systems Cost Analysis Group, 22-23 October 2002, 37 charts.
- S.A. Book and E.L. Burgess, "The Learning Rate's Overpowering Impact on Cost Estimates and How to Diminish It", *Journal of Parametrics*, Vol. 16 (1996), pages 33-57.
- S.A. Book and N.Y. Lao, "Minimum-Percentage-Error Regression under Zero-Bias Constraints", *Proceedings of the Fourth Annual U.S. Army Conference on Applied Statistics, 21-23 October 1998*, U.S. Army Research Laboratory, Report No. ARL-SR-84, November 1999, pages 47-56.
- S.A. Book and J.C. Latta, "'Learning' Vs. 'Estimating': Issues In Modeling Cost Improvement", The Aerospace Corporation, Unclassified Report Prepared for the NRO Cost Group, December 2000, 9 pages.
- J.C. Latta, "Composite Cost Improvement," Space Systems Cost Analysis Group, 22-23 October 2002, 23 charts.
- Lee, D.A., *The Cost Analyst's Companion*, McLean, VA: Logistics Management Institute, 1997, 139+xi pages.
- D. MacKenzie and B. Addison, "Space Hardware Cost Improvement Curves," ISPA 1999 International Conference, 26 charts.

An AT&T Company

Glossary

- **AUC = Average Unit Cost**
- **CER = Cost Estimating Relationship**
- ***log* = Natural *or* Common Logarithm (as long as usage is consistent)**
- **LR = Learning Rate (as a decimal)**
- ***N* = Lot Size (Number of Units in Lot)**
- ***Q* = Prior Quantity**
- **QAIV = Quantity as an Independent Variable**
- **R^2 = Square of Pearson's Correlation Between Estimates and Data-Base "Actuals"**
- ***W* = Weight**

Author Bios

Dr. Stephen A. Book is Chief Technical Director of MCR, Inc. In that capacity, he is responsible for ensuring technical excellence of MCR products, services, and processes by encouraging process improvement, maintaining quality control, and training employees and customers in cost and schedule analysis and associated program-control disciplines. Dr. Book joined MCR in January 2001 after 21 years with The Aerospace Corporation, holding the title "Distinguished Engineer" during 1996-2000 and having served as Director, Resource and Requirements Analysis Department, during 1989-1995. While at The Aerospace Corporation, he directed a vigorous program of cost research into methods of conducting cost and schedule risk analyses and deriving cost-estimating relationships. Dr. Book has given numerous technical and tutorial presentations on cost-risk analysis and other statistical aspects of costing to DoD, NASA, and ESA Cost Symposia, the AF/NASA/ESA Space Systems Cost Analysis Group (SSCAG), and various professional societies. He has served on national panels reviewing NASA programs, such as the 1997-98 Cost Assessment and Validation Task Force on the International Space Station and the 1998-99 National Research Council Committee on Space Shuttle Upgrades. He is the current chair of the Risk Subgroup of SSCAG and a member of the Economics Technical Committee of the American Institute of Aeronautics and Astronautics (AIAA). Dr. Book earned his Ph.D. in mathematics, with concentration in probability and statistics, at the University of Oregon.

Mr. Erik L. Burgess is Technical Manager at MCR, Inc., with primary responsibility for providing cost and budget-profiling analysis to the National Reconnaissance Office (NRO) Cost Group. He also conducts technical quality reviews on MCR's efforts in support of various DoD and other government agencies. Mr. Burgess joined MCR in March 2002 after three years at PricewaterhouseCoopers, LLP, Fairfax, VA, and eight years at The Aerospace Corporation. While at PricewaterhouseCoopers, Mr. Burgess served as technical manager for the Navy's Visibility and Management of Operating and Support Costs (VAMOSC) "datawarehouse" and provided supply-chain modeling support to the ADUSD (Logistics Plans and Programs) during the 2001 Quadrennial Defense Review. While at The Aerospace Corporation, he conducted engineering analysis, cost research, and cost estimating on several concept studies for program offices at the USAF Space and Missile Systems Center, the NRO, and NASA. He also served as Senior Project Engineer on a satellite program during its concept exploration and pre-acquisition phases. Mr. Burgess earned his M.S. and B.S. in Mechanical Engineering from the Massachusetts Institute of Technology.